

SEQUENCE LISTING

<110> CHEN, WENFANG
MEEK, THOMAS D.
POWELL, DAVID J.
TEW, DAVID G.

<120> Method of Site Specific Labeling of Proteins and Uses
Therefor

<130> P50892

<140> 09/889,344
<141> 2001-07-16

<150> PCT/US00/01481
<151> 2000-01-20

<150> US 60/117,327
<151> 1999-01-22

<160> 16

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<221> unsure
<222> (5)
<223> Where Xaa at position (5) can represent Leucine or Isoleucine

<400> 1

Gln Ser Lys Val Xaa

1

5

<210> 2
<211> 207
<212> PRT
<213> Artificial Sequence

<220>
<221> unsure
<222> (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15)
<222> (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28)
<222> (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41)
<222> (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54)
<222> (55) (56) (57) (58) (59) (60) (61) (62) (62) (64) (65) (66) (67)
<222> (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80)
<222> (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93)
<222> (94) (95) (96) (97) (98) (99) (100) (101) (106) (107) (108) (109)
<222> (110) (111) (112) (113) (114) (115) (116) (117) (118) (119) (120)
<222> (121) (122) (123) (124) (125) (126) (127) (128) (129) (130) (131)
<222> (132) (133) (134) (135) (136) (137) (138) (139) (140) (141) (142)
<222> (143) (144) (145) (146) (147) (148) (149) (150) (151) (152) (153)
<222> (154) (155) (156) (157) (158) (159) (160) (161) (162) (163) (164)
<222> (165) (166) (167) (168) (169) (170) (171) (172) (173) (174) (175)
<222> (176) (177) (178) (179) (180) (181) (182) (183) (184) (185) (186)
<222> (187) (188) (189) (190) (191) (192) (193) (194) (195) (196) (197)
<222> (198) (199) (200) (201) (202) (203) (204) (205) (206) (207)
<223> Where Xaa can represent none or any one of the twenty naturally
occurring amino acids

<400> 2

Xaa Xaa

1

5

10

15

Xaa Xaa

20

25

30

Xaa Xaa

35

40

45

Xaa Xaa

50

55

60

Xaa
65 70 75 80
Xaa
85 90 95
Xaa Xaa Xaa Xaa Xaa Gln Ser Lys Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110
Xaa
115 120 125
Xaa
130 135 140
Xaa
145 150 155 160
Xaa
165 170 175
Xaa
180 185 190
Xaa
195 200 205

<210> 3

<211> 207

<212> PRT

<213> Artificial Sequence

<220>

<221> unsure

<222> (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15)

<222> (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28)

<222> (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41)

<222> (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54)

<222> (55) (56) (57) (58) (59) (60) (61) (62) (62) (64) (65) (66) (67)

<222> (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80)

<222> (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93)

<222> (94) (95) (96) (97) (98) (99) (100) (101) (106) (107) (108) (109)

<222> (110) (111) (112) (113) (114) (115) (116) (117) (118) (119) (120)

<222> (121) (122) (123) (124) (125) (126) (127) (128) (129) (130) (131)

<222> (132) (133) (134) (135) (136) (137) (138) (139) (140) (141) (142)

<222> (143) (144) (145) (146) (147) (148) (149) (150) (151) (152) (153)

<222> (154) (155) (156) (157) (158) (159) (160) (161) (162) (163) (164)

<222> (165) (166) (167) (168) (169) (170) (171) (172) (173) (174) (175)

<222> (176) (177) (178) (179) (180) (181) (182) (183) (184) (185) (186)
<222> (187) (188) (189) (190) (191) (192) (193) (194) (195) (196) (197)
<222> (198) (199) (200) (201) (202) (203) (204) (205) (206) (207)
<223> Where Xaa can represent none or any one of the twenty naturally
occurring amino acids

<400> 3

Xaa
1 5 10 15
Xaa
20 25 30
Xaa
35 40 45
Xaa
50 55 60
Xaa
65 70 75 80
Xaa
85 90 95
Xaa Xaa Xaa Xaa Gln Ser Lys Val Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110
Xaa
115 120 125
Xaa
130 135 140
Xaa
145 150 155 160
Xaa
165 170 175
Xaa
180 185 190
Xaa
195 200 205

<210> 4

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Derivative of a factor XIII substrate

<400> 4

Leu Ser Leu Ser Gln Ser Lys Val Leu Gly

1

5

10

<210> 5

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Derivative of a factor XIII substrate

<400> 5

Ile Gly Glu Gly Gln Ser Lys Val Leu Gly

1

5

10

<210> 6

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Derivative of a factor XIII substrate

<400> 6

Leu Gly Pro Gly Gln Ser Lys Val Ile Gly

1

5

10

<210> 7

<211> 81

<212> DNA

<213> Unknown

<220>

<223> Oligonucleotide designed to introduce Q tag

<400> 7

tgtacctcag accatatgag cctgtccctg tcccagtcca aagttctgcc gggtccgagc

60

actatcgaag aacgcgttaa g

81

<210> 8
<211> 37
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide designed to introduce Q tag

<400> 8

tgatgtcagt caagcttacg cctgggtggcc gttgatg

37

<210> 9
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Derivative of a factor XIII substrate

<400> 9

Met Ser Leu Ser Leu Ser Gln Ser Lys Val Leu Pro Gly Pro
1 5 10

<210> 10
<211> 37
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide designed to introduce Q tag

<400> 10

tgtacaccttag accatatgag cactatcgaa gaacgcg

37

<210> 11
<211> 78
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide designed to introduce Q tag

<400> 11
tgatgtcagt caagcttacg gaccggcag aactttggac tgggacaggg acagcgccctg 60
gtggccgttg atgtaatc 78

<210> 12
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> Derivative of E. coli ACP protein

<400> 12
Leu Ser Leu Ser Gln Ser Lys Val Leu Pro Gly Pro
1 5 10

<210> 13
<211> 92
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide designed to introduce Q tag into
Streptococcus haemophilus FabH gene

<400> 13
tatcatatga gcctgtccct gtcccagtcc aaagttctgc cgggtccggg taccctcgag 60
ggatccgctt ttgcaaaaat aagtcaaggaa gc 92

<210> 14
<211> 53
<212> DNA
<213> Unknown

<220>
<223> Oligonucleotide designed to introduce Q tag into

Streptococcus haemophilus FabH gene

<400> 14
ctcagatctg agtcactag tggatcctta aattgttata atgagcggtgc ccc 53

<210> 15
<211> 364
<212> PRT
<213> Artificial Sequence

<220>
<223> Modified sequence of Streptococcus haemophilus FabH

<400> 15

Met	Gly	His	Ser	Ser	Gly	His									
1		5				10				15					
Ile	Glu	Gly	Arg	His	Met	Ser	Leu	Ser	Leu	Ser	Gln	Ser	Lys	Val	Leu
	20					25				30					
Pro	Gly	Pro	Gly	Thr	Leu	Glu	Gly	Ser	Ala	Phe	Ala	Lys	Ile	Ser	Gln
	35				40				45						
Val	Ala	His	Tyr	Val	Pro	Glu	Gln	Val	Val	Thr	Asn	His	Asp	Leu	Ala
	50				55			60							
Gln	Ile	Met	Asp	Thr	Asn	Asp	Glu	Trp	Ile	Ser	Ser	Arg	Thr	Gly	Ile
	65				70			75			80				
Arg	Gln	Arg	His	Ile	Ser	Arg	Thr	Glu	Ser	Thr	Ser	Asp	Leu	Ala	Thr
	85				90			95							
Glu	Val	Ala	Lys	Lys	Leu	Met	Ala	Lys	Ala	Gly	Ile	Thr	Gly	Lys	Glu
	100				105			110							
Leu	Asp	Phe	Ile	Ile	Leu	Ala	Thr	Ile	Thr	Pro	Asp	Ser	Met	Met	Pro
	115				120			125							
Ser	Thr	Ala	Ala	Arg	Val	Gln	Ala	Asn	Ile	Gly	Ala	Asn	Lys	Ala	Phe
	130				135			140							
Ala	Phe	Asp	Leu	Thr	Ala	Ala	Cys	Ser	Gly	Phe	Val	Phe	Ala	Leu	Ser
	145				150			155			160				
Thr	Ala	Glu	Lys	Phe	Ile	Ala	Ser	Gly	Arg	Phe	Gln	Lys	Gly	Leu	Val
	165				170			175							
Ile	Gly	Ser	Glu	Thr	Leu	Ser	Lys	Ala	Val	Asp	Trp	Ser	Asp	Arg	Ser
	180				185			190							
Thr	Ala	Val	Leu	Phe	Gly	Asp	Gly	Ala	Gly	Gly	Val	Leu	Glu	Ala	
	195				200			205							

Ser Glu Gln Glu His Phe Leu Ala Glu Ser Leu Asn Ser Asp Gly Ser
210 215 220
Arg Ser Glu Cys Leu Thr Tyr Gly His Ser Gly Leu His Ser Pro Phe
225 230 235 240
Ser Asp Gln Glu Ser Ala Asp Ser Phe Leu Lys Met Asp Gly Arg Thr
245 250 255
Val Phe Asp Phe Ala Ile Arg Asp Val Ala Lys Ser Ile Lys Gln Thr
260 265 270
Ile Asp Glu Ser Pro Ile Glu Val Thr Asp Leu Asp Tyr Leu Leu Leu
275 280 285
His Gln Ala Asn Asp Arg Ile Leu Asp Lys Met Ala Arg Lys Ile Gly
290 295 300
Val Asp Arg Ala Lys Leu Pro Ala Asn Met Met Glu Tyr Gly Asn Thr
305 310 315 320
Ser Ala Ala Ser Ile Pro Ile Leu Leu Ser Glu Cys Val Glu Gln Gly
325 330 335
Leu Ile Pro Leu Asp Gly Ser Gln Thr Val Leu Leu Ser Gly Phe Gly
340 345 350
Gly Gly Leu Thr Trp Gly Thr Leu Ile Leu Thr Ile
355 360

<210> 16

<211> 503

<212> PRT

<213> Artificial Sequence

<220>

<223> Modified sequence of Erythropoietin receptor
fusion protein

<400> 16

Met Asp His Leu Gly Ala Ser Leu Trp Pro Gln Val Gly Ser Leu Cys
1 5 10 15
Leu Leu Leu Ala Gly Ala Ala Trp Ala Pro Pro Pro Asn Leu Pro Asp
20 25 30
Pro Lys Phe Glu Ser Lys Ala Ala Leu Leu Ala Ala Arg Gly Pro Glu
35 40 45
Glu Leu Leu Cys Phe Thr Glu Arg Leu Glu Asp Leu Val Cys Phe Trp
50 55 60

Glu Glu Ala Ala Ser Ala Gly Val Gly Pro Gly Asn Tyr Ser Phe Ser
65 70 75 80
Tyr Gln Leu Glu Asp Glu Pro Trp Lys Leu Cys Arg Leu His Gln Ala
85 90 95
Pro Thr Ala Arg Gly Ala Val Arg Phe Trp Cys Ser Leu Pro Thr Ala
100 105 110
Asp Thr Ser Ser Phe Val Pro Leu Glu Leu Arg Val Thr Ala Ala Ser
115 120 125
Gly Ala Pro Arg Tyr His Arg Val Ile His Ile Asn Glu Val Val Leu
130 135 140
Leu Asp Ala Pro Val Gly Leu Val Ala Arg Leu Ala Asp Glu Ser Gly
145 150 155 160
His Val Val Leu Arg Trp Leu Pro Pro Glu Thr Pro Met Thr Ser
165 170 175
His Ile Arg Tyr Glu Val Asp Val Ser Ala Gly Asn Gly Ala Gly Ser
180 185 190
Val Gln Arg Val Glu Ile Leu Glu Gly Arg Thr Glu Cys Val Leu Ser
195 200 205
Asn Leu Arg Gly Arg Thr Arg Tyr Thr Phe Ala Val Arg Ala Arg Met
210 215 220
Ala Glu Pro Ser Phe Gly Gly Phe Trp Ser Ala Trp Ser Glu Pro Val
225 230 235 240
Ser Leu Leu Thr Pro Ser Asp Leu Asp Pro Leu Ser Leu Ser Gln Ser
245 250 255
Lys Val Leu Gly Val Phe Phe Ala Glu Ile Glu Gly Arg Gly Thr Glu
260 265 270
Pro Lys Ser Ala Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
275 280 285
Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
290 295 300
Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
305 310 315 320
Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp
325 330 335
Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr
340 345 350
Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
355 360 365
Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu
370 375 380

Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
385 390 395 400
Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys
405 410 415
Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
420 425 430
Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
435 440 445
Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
450 455 460
Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser
465 470 475 480
Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
485 490 495
Leu Ser Leu Ser Pro Gly Lys
500